



# On the Construction of Group Equivariant Non-Expansive Operators *via* Permutants and Symmetric Functions

Francesco Conti<sup>1,2</sup>, Patrizio Frosini<sup>3,4,5,6\*</sup> and Nicola Quercioli<sup>3,7</sup>

<sup>1</sup> Department of Mathematics, University of Pisa, Pisa, Italy, <sup>2</sup> Institute of Information Science and Technologies "A. Faedo", National Research Council of Italy (CNR), Pisa, Italy, <sup>3</sup> Department of Mathematics, University of Bologna, Bologna, Italy, <sup>4</sup> Alma Mater Research Center on Applied Mathematics, University of Bologna, Bologna, Italy, <sup>5</sup> Alma Mater Research Institute for Human-Centered Artificial Intelligence, University of Bologna, Bologna, Italy, <sup>6</sup> Research Centre on Electronic Systems for the Information and Communication Technology, University of Bologna, Bologna, Italy, <sup>7</sup> ENEA Centro Ricerche Bologna, Bologna, Italy

## OPEN ACCESS

### Edited by:

Fabio Anselmi,  
Baylor College of Medicine,  
United States

### Reviewed by:

Remco Duits,  
Eindhoven University of Technology,  
Netherlands  
Kelin Xia,  
Nanyang Technological University,  
Singapore

### \*Correspondence:

Patrizio Frosini  
patrizio.frosini@unibo.it

### Specialty section:

This article was submitted to  
Machine Learning and Artificial  
Intelligence,  
a section of the journal  
Frontiers in Artificial Intelligence

**Received:** 29 September 2021

**Accepted:** 18 January 2022

**Published:** 15 February 2022

### Citation:

Conti F, Frosini P and Quercioli N  
(2022) On the Construction of Group  
Equivariant Non-Expansive Operators  
*via* Permutants and Symmetric  
Functions.  
*Front. Artif. Intell.* 5:786091.  
doi: 10.3389/frai.2022.786091

Group Equivariant Operators (GEOs) are a fundamental tool in the research on neural networks, since they make available a new kind of geometric knowledge engineering for deep learning, which can exploit symmetries in artificial intelligence and reduce the number of parameters required in the learning process. In this paper we introduce a new method to build non-linear GEOs and non-linear Group Equivariant Non-Expansive Operators (GENEOs), based on the concepts of symmetric function and permutant. This method is particularly interesting because of the good theoretical properties of GENEOs and the ease of use of permutants to build equivariant operators, compared to the direct use of the equivariance groups we are interested in. In our paper, we prove that the technique we propose works for any symmetric function, and benefits from the approximability of continuous symmetric functions by symmetric polynomials. A possible use in Topological Data Analysis of the GENEOs obtained by this new method is illustrated.

**Keywords:** GENE, permutant, symmetric function, persistence diagram, persistent homology, machine learning

## 1. INTRODUCTION

In recent years, the theory of equivariant operators has become a topic of great interest to the scientific community, since these operators allow to make explicit the use of symmetries in deep learning and artificial intelligence (Mallat, 2012, 2016; Bengio et al., 2013; Zhang et al., 2015; Anselmi et al., 2016, 2019; Cohen and Welling, 2016; Worrall et al., 2017), thereby reducing the number of parameters required in the learning process. In particular, group equivariant non-expansive operators (GENEOs) have been recently proposed as elementary components for building new kinds of neural networks, benefiting from good mathematical properties, such as compactness and convexity, under suitable assumptions on the space of data and with respect to the choice of appropriate topologies (Bergomi et al., 2019). In particular, compactness guarantees total-boundedness, i.e., for every  $\varepsilon > 0$  we can find a finite set  $\mathcal{F} = \{F_1, \dots, F_s\}$  of GENEOs such that any other GENEO has a distance less than  $\varepsilon$  from at least one  $F_i$ . This property opens the way to the search for methods to effectively build such a representative set  $\mathcal{F}$ , leading us to look for new techniques to produce GENEOs.

GENEOs are grounded in Topological Data Analysis (TDA) and allow to shift the attention from the data to the observers who process them, and to the properties of invariance and simplification associated with those observers. The use of these operators is justified by the fact that in most of the cases we are not directly interested in data, but in approximating the experts' behavior in the presence of given data (Frosini, 2016). Since different agents can have different reactions in the presence of the same data, it is clear that data analysis has to be based on each pair (data, observer) rather than on data alone. From the point of view of AI, the focus on GENEOS corresponds to the rising interest in the so called "explainable deep learning" (Rudin, 2019; Carrieri et al., 2021; Hicks et al., 2021), which looks for methods and techniques that can be understood by humans.

GENEOs transform data according to two properties. First of all, they are equivariant with respect to the action of a given transformation group, i.e., they commute with such a group. Secondly, they do not increase the distance between data. This kind of regularity is frequently found in applications, since in several cases the operators we use are required to simplify the metric structure of data. We can obviously imagine particular applications where this condition is locally violated, but the usual long-term goal is to produce representations that are much simpler and more meaningful than the original data, thereby leading us to assume that the considered (compositions of any sufficiently long chain of) operators are non-expansive. This assumption is not only of use to simplify the information we have to manage, but it is also fundamental in the proof that the space of group equivariant non-expansive operators is compact (and hence finitely approximable), provided that the space of data is compact with respect to a suitable topology (Bergomi et al., 2019). This statement becomes false if we renounce non-expansivity.

The use of GENEOS is not limited to machine learning. Another important reason for the study of these operators follows from the relationship between GENEOS and TDA. We indeed know that TDA and Persistent Homology allow for a qualitative and efficient geometric study of the data space, but suffer from some important limitations, since Persistent Homology alone is not able to distinguish between some functions. Fortunately, the joint use of TDA and GENEOS overcomes this difficulty in the discrimination of data (Frosini, 2016; Frosini and Jabłoński, 2016; Bergomi et al., 2019). In other words, GENEOS are able to preserve information on the data that would have been lost through TDA alone.

Another interesting aspect of GENEOS is that we can also look at them as operators that change the pseudo-metrics we use in data comparison. If the real-valued functions  $\varphi_1, \varphi_2$  represent the data we have to compare and a GENEOS  $F$  is given, we can replace the max-norm distance  $\|\varphi_1 - \varphi_2\|_\infty$  with the new pseudo-metric  $d'(\varphi_1, \varphi_2) := \|F(\varphi_1) - F(\varphi_2)\|_\infty$ . In this approach,  $F$  is not seen as a map that transforms the data we are considering, but as a new way of comparing data. We will see in section 6 that the availability of non-linear GENEOS can indeed produce more flexible pseudo-metrics.

Last but not least, a theory of GENEOS could be a relevant tool in the investigation of the role of internal conflicts in AI. We know that the availability of procedures that emulate intelligence

opens the way to the appearance of contradictions, conflicts and unexpected behaviors (Frosini, 2009). This phenomenon cannot be ignored in the mathematical study of AI. The use of a precise geometric formalization of components in machine learning could be of great help in facing and analyzing this emerging problem.

However, the main reason for the research about GENEOS follows from a shift of interest from the spaces of data to the topological and geometric analysis of the spaces of observers of the data. This fact naturally leads us to the problem of the efficient approximation of observers. Such an approximation requires to make available large and dense sets of GENEOS, each one representing a possible data-observer interaction. Therefore, since non-linear interactions between observers and data are of great importance in applications, new techniques to build non-linear GENEOS are needed. The main contribution of this paper consists in introducing a new method to produce non-linear GENEOS through the concepts of symmetric function and permutant, thereby extending the procedure illustrated in Botteghi et al. (2020) for the building of linear GENEOS. In this way, we strictly expand the set of operators we can use in applications.

The concept of permutant comes into play when a set  $\Phi$  of functions from a space  $X$  to  $\mathbb{R}$  and a group  $G$  of permutations of  $X$  are given. The set  $\Phi$  represents the space of signals we are interested in, and is assumed to be preserved by right composition with elements of  $G$ . If two signals  $\varphi_1, \varphi_2 \in \Phi$  are obtained from each other by right composition with an element  $g \in G$ , we say that they are equivalent with respect to  $G$ , just as happens when two images are considered equivalent if there exists an isometry changing one into the other. In this setting, a *permutant* is defined as a finite set  $H$  of  $\Phi$ -preserving permutations of  $X$  that is stable under the conjugation action  $h \mapsto g \circ h \circ g^{-1}$  of any element of  $G$  on  $H$  (Camporesi et al., 2018).

This paper shows that when a symmetric function and a permutant for the equivariance group  $G$  are available, we can easily build a (non-linear) GENEOS with respect to  $G$  (section 3). This fact justifies the theoretical and practical importance of permutants. Our long-term purpose is the one of developing an effective theory for the approximation of observers and agents *via* GENEOS in a topological-geometrical setting, so extending the use of these operators in deep learning. While this goal is challenging, we think that our approach could lead to think of GENEOS as elementary components in the building of a new kind of neural networks. This idea is justified by at least two reasons. First of all, deep learning could benefit from using components that are guaranteed to be equivariant with respect to given groups of transformations and are grounded in a well founded topological theory, thereby allowing neural nets to save time in the learning process and to take advantage of techniques developed in TDA. Secondly, an engineering based on GENEOS would be much more transparent, because of the intrinsic interpretability of its components.

The reader could wonder why building GENEOS *via* permutants should be better than building them by other methods (for example by integrating on the equivariance group

G). The key point is that in many applications some permutants exist, whose size is much smaller than the size of the equivariance group. In these cases, the approaches based on permutants can be much simpler than the ones based on  $G$ . We observe that permutants encode part of the information represented by the data equivalence expressed by  $G$ . Of course, by deciding to build GENEOS *via* permutants we implicitly accept to lose some information about such a data equivalence, and make a compromise between the computational complexity and the analytical power of the operators we are interested in. The reader can understand this tradeoff by thinking about the limit case given by a permutant containing just the identical permutation  $\text{id}$  of  $X$ . While the singleton  $\{\text{id}\}$  is indeed a (trivial) permutant, it does not give any information about the equivariance group  $G$  we are considering, since  $\{\text{id}\}$  is a permutant for *any* group of  $\Phi$ -preserving permutations of  $X$ . However, if we consider a larger and larger set  $H$  of  $\Phi$ -preserving permutations of  $X$ , the set of groups admitting  $H$  as a permutant becomes smaller and smaller. In other words, larger permutants make easier the identification of  $G$ .

This article is part of an extensive research on permutants. In Botteghi et al. (2020) it has been proved that each linear  $G$ -equivariant non-expansive operator can be produced by a weighted summation associated with a suitable weighted permutant, provided that the group  $G$  transitively acts on a finite signal domain. This paper opens the way to the research about the natural conjecture that each non-linear  $G$ -equivariant non-expansive operator can be produced (or at least well approximated) by applying our new technique to suitable symmetric functions and permutants, provided that the group  $G$  transitively acts on a finite signal domain. This probably non-trivial problem will be attacked in following papers, grounding on the results obtained in this article.

The outline of the paper is as follows. In section 2, we recall the main definitions in our mathematical setting. In Section 3, we show how to associate a group equivariant operator (GEO) with a symmetric function. Section 4 is devoted to the approximation of a generic continuous symmetric function by a polynomial in the elementary symmetric functions, and in section 5, we finally show how to associate a GENEOS with such a polynomial. Section 6 highlights the benefits of our approach.

For more details and proofs about the results and concepts illustrated in section 2 we refer the interested reader to the papers (Frosini, 2016; Frosini and Jabłoński, 2016; Frosini and Quercioli, 2017; Camporesi et al., 2018; Bergomi et al., 2019). The other sections present our new results about the construction of non-linear GENEOS *via* symmetric functions and permutants.

## 2. MATHEMATICAL SETTING

Let  $X$  be a non-empty set and consider a non-empty, compact subspace  $\Phi$  of the normed vector space  $(\mathbb{R}_b^X, \|\cdot\|_\infty)$ , where  $\mathbb{R}_b^X$  is the set of all bounded real-valued functions with domain  $X$ , and  $\|\varphi\|_\infty := \sup_{x \in X} |\varphi(x)|$ . We can think of the functions in  $\Phi$  as the data, i.e., the measurements provided by our measuring instruments (or by any operator), and of  $X$  as the space where

the measurements are made. Sometimes the functions in  $\Phi$  are also referred as **admissible filtering functions** or **admissible signals**. We now recall the usual setting for the introduction of group equivariant non-expansive operators. We endow  $\Phi$  with the topology induced by the uniform convergence distance

$$D_\Phi(\varphi_1, \varphi_2) := \|\varphi_1 - \varphi_2\|_\infty.$$

At this stage,  $X$  is only a set. We endow  $X$  with the topology induced by the pseudo-metric

$$D_X(x_1, x_2) := \sup_{\varphi \in \Phi} |\varphi(x_1) - \varphi(x_2)|.$$

The idea behind this definition is that two points  $x_1, x_2 \in X$  are considered different only if they are taken to different values by at least one admissible filtering function.

We recall that a pseudo-metric space is a generalization of a metric space in which the distance between two distinct points can be zero. Moreover, a function  $f$  from a pseudo-metric space  $(P_1, d_1)$  to a pseudo-metric space  $(P_2, d_2)$  is called **non-expansive** if

$$d_2(f(x), f(y)) \leq d_1(x, y)$$

for every  $x, y \in P_1$ .

**Remark 2.1.** Every function  $\varphi \in \Phi$  is non-expansive with respect to the pseudo-metric  $D_X$  on  $X$  and the Euclidean metric on  $\mathbb{R}$ . Therefore, each function  $\varphi \in \Phi$  is continuous with respect to these topologies.

Since  $\Phi$  is compact, the topology induced by the pseudo-metric  $D_X$  coincides with the initial topology  $\tau_{in}$  on  $X$  with respect to  $\Phi$  (see Theorem 2.1 in Bergomi et al., 2019, Supplementary Methods). We recall that the initial topology is the coarsest topology on  $X$  which makes each function in  $\Phi$  continuous. Moreover, the compactness of  $\Phi$  implies that if  $X$  is complete then it is also compact (see Theorem 2.2 in Bergomi et al., 2019, Supplementary Methods). In this work, we assume that  $X$  is complete, and therefore compact with respect to the topology induced by  $D_X$ . The image of  $X$  through the filtering functions is denoted by  $\text{Im}(\Phi)$  and is defined as

$$\text{Im}(\Phi) = \{\varphi(x) \text{ s.t. } \varphi \in \Phi, x \in X\}.$$

The following result will be of use in section 5.

**Proposition 2.2.** If  $X$  and  $\Phi$  are compact,  $\text{Im}(\Phi)$  is compact with respect to the Euclidean topology.

*Proof:* Let us consider the function  $\gamma: \Phi \times X \rightarrow \mathbb{R}$  such that  $\gamma(\varphi, x) := \varphi(x)$ . The space  $\Phi \times X$  is compact with respect to the product topology, which we recall is induced by the sum pseudo-distance. Since the continuous image of a compact is compact,  $\gamma(\Phi \times X) = \text{Im}(\Phi)$ , and every non-expansive function is continuous, it is sufficient to prove that  $\gamma$  is non-expansive. Given  $\varphi_1, \varphi_2 \in \Phi$  and  $x_1, x_2 \in X$ , we have that

$$|\gamma(\varphi_1, x_1) - \gamma(\varphi_2, x_2)| = |\varphi_1(x_1) - \varphi_2(x_2)|$$

$$\begin{aligned}
&= |\varphi_1(x_1) - \varphi_1(x_2) + \varphi_1(x_2) - \varphi_2(x_2)| \\
&\leq \sup_{\varphi \in \Phi} |\varphi(x_1) - \varphi(x_2)| \\
&+ \sup_{x \in X} |\varphi_1(x) - \varphi_2(x)| \\
&= D_X(x_1, x_2) + D_\Phi(\varphi_1, \varphi_2).
\end{aligned}$$

We have proved that  $\gamma$  is non-expansive, and therefore  $\text{Im}(\Phi)$  is compact.

**Definition 2.3.** Chachólski et al. (2020) A  $\Phi$ -operation is a function  $g: X \rightarrow X$  such that, for every  $\varphi \in \Phi$ , the composition  $\varphi \circ g$  also belongs to  $\Phi$ .

**Definition 2.4.** A  $\Phi$ -operation  $g$  is *invertible* if there is a  $\Phi$ -operation  $h$  such that  $g \circ h = h \circ g = \text{id}_X$ .

We denote the collection of all invertible  $\Phi$ -operations by  $\text{Aut}_\Phi(X)$ . In other words,

$$\begin{aligned}
\text{Aut}_\Phi(X) &= \{g: X \rightarrow X \mid g \text{ is a bijection, and } \varphi \circ g, \\
&\varphi \circ g^{-1} \in \Phi \text{ for every } \varphi \in \Phi\}.
\end{aligned}$$

We note that  $\text{Aut}_\Phi(X)$  is a group with respect to the usual composition operation.

**Definition 2.5.** A *perception pair* is an ordered pair  $(\Phi, G)$  where  $\Phi \subseteq \mathbb{R}_b^X$  and  $G$  is a subgroup of  $\text{Aut}_\Phi(X)$ .

As an example,  $(\Phi, \text{Aut}_\Phi(X))$  is always a perception pair.

**Remark 2.6.** When a perception pair  $(\Phi, G)$  is given, each element  $g \in G$  acts on the set  $\Phi$  by right composition, taking each function  $\varphi \in \Phi$  to the function  $\varphi \circ g$ .

## 2.1. Group Equivariant Non-Expansive Operators

**Definition 2.7.** Let us consider two perception pairs  $(\Phi, G)$ ,  $(\Psi, H)$  and a homomorphism  $T: G \rightarrow H$ . Each map  $F: \Phi \rightarrow \Psi$  such that  $F$  is  $T$ -equivariant (i.e.,  $F(\varphi \circ g) = F(\varphi) \circ T(g)$  for every  $\varphi \in \Phi, g \in G$ ) is called a **Group Equivariant Operator (GEO)** with respect to  $T$ .

**Definition 2.8.** Let us consider two perception pairs  $(\Phi, G)$ ,  $(\Psi, H)$  and a homomorphism  $T: G \rightarrow H$ . Each map  $F: \Phi \rightarrow \Psi$  such that  $F$  is  $T$ -equivariant (i.e.,  $F(\varphi \circ g) = F(\varphi) \circ T(g)$  for every  $\varphi \in \Phi, g \in G$ ) and non-expansive (i.e.,  $\|F(\varphi) - F(\psi)\|_\infty \leq \|\varphi - \psi\|_\infty$  for every  $\varphi, \psi \in \Phi$ ) is called a **Group Equivariant Non-Expansive Operator (GENEO)** with respect to  $T$ .

After fixing two perception pairs  $(\Phi, G)$ ,  $(\Psi, H)$  and a homomorphism  $T: G \rightarrow H$ , we will use the symbol  $\mathcal{F}_T^{\text{all}}$  to denote the collection of all GENEOS with respect to  $T$  between such perception pairs. We endow  $\mathcal{F}_T^{\text{all}}$  with the topology induced by the metric  $D_{\text{GENEO}}(F_1, F_2) := \sup_{\varphi \in \Phi} D_\Phi(F_1(\varphi), F_2(\varphi))$ . For a more in-depth study of the GENEOS topology, we refer the reader to Bergomi et al. (2019). We stress that the non-expansivity of the operators is pivotal for two reasons. The first

reason is that we want our operators to simplify the data metric, i.e., not to introduce complexity into the data. The second reason is that non-expansivity allows us to prove the compactness of the space  $\mathcal{F}_T^{\text{all}}$ , provided that  $\Phi$  and  $\Psi$  are compact with respect to the distances  $D_\Phi, D_\Psi$  (see Theorem 7 in Bergomi et al., 2019). If we remove the assumption that our operators are non-expansive, this property of compactness does not hold anymore. As an example, let  $\Phi = \Psi$  be equal to the set of all constant functions from  $\mathbb{R}$  to  $[0, 1]$ , and  $G = H$  be the trivial group containing just the identity permutation of  $\mathbb{R}$ . We observe that  $\Phi, X = \mathbb{R}$  and  $G$  are compact with respect to the topologies we have defined on them. Let us now consider the sequence  $(F_n)$  of GEOs from  $\Phi$  to  $\Phi$  with respect to the identity homomorphism  $\text{id}_G: G \rightarrow G$ , defined by setting  $F_n(\varphi) := \varphi^n$  for every function  $\varphi \in \Phi$  and every positive integer  $n$ . It is easy to check that  $\lim_{n \rightarrow \infty} D_{\text{GENEO}}(F_m, F_n) = 1$  for every positive integer  $m$ , and hence the sequence  $(F_n)$  does not admit any converging subsequence. This implies that the space of all GEOs from  $\Phi$  to  $\Phi$  with respect to  $\text{id}_G$  is not compact. The compactness of  $\mathcal{F}_T^{\text{all}}$  is a key property in applications, since it guarantees that such a space can be approximated by a finite set.

If  $G = H$  and  $T = \text{id}_G$ , we can say that  $F: \Phi \rightarrow \mathbb{R}_b^X$  is a  $G$ -equivariant map. From now on, we will make these assumptions, and use the terms GEO and GENEOS with reference to this setting.

## 2.2. Permutants

**Definition 2.9.** Let  $S_X$  be the set of permutations of  $X$ . For each  $g \in G$ , the map  $c_g: S_X \rightarrow S_X$  taking each  $s \in S_X$  to  $g \circ s \circ g^{-1}$  is called the **conjugation action** of  $g \in G$  on  $S_X$ . For every subset  $H$  of  $S_X$ , we denote the set  $c_g(H)$  by the symbol  $gHg^{-1}$ .

**Definition 2.10.** Camporesi et al. (2018) A finite set  $H \subseteq \text{Aut}_\Phi(X)$  is called a **permutant** for  $G$  if either  $H = \emptyset$  or  $gHg^{-1} = H$  for every  $g \in G$ .

**Remark 2.11.** In general, a permutant is not a normal subgroup of  $G$ . Indeed we require neither that  $H$  is a group nor that  $H$  is a subset of  $G$ . We observe that the sets  $\emptyset$  and  $\{\text{id}_X\}$  are trivial permutants for any subgroup  $G$  of  $\text{Aut}_\Phi(X)$ . Both  $G$  and  $\text{Aut}_\Phi(X)$  are also permutants for  $G$ , provided that they are finite groups.

**Example 2.12.** Let  $\Phi$  be the set of all functions  $\varphi: X = S^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\} \rightarrow [0, 1]$  that are non-expansive with respect to the Euclidean distances on  $S^1$  and  $[0, 1]$ . Let us consider the group  $G$  of all isometries of  $\mathbb{R}$ . If  $h$  is the clockwise rotation of  $\ell$  radians for a fixed  $\ell \in \mathbb{R}$ , then the set  $H = \{h, h^{-1}\}$  is a permutant for  $G$ .

**Example 2.13.** Let  $\Phi$  be the set of all functions  $\varphi: X = S^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\} \rightarrow [0, 1]$  that are non-expansive with respect to the Euclidean distances on  $S^1$  and  $[0, 1]$ . Let  $G$  be the group generated by the reflection with respect to the axis  $x = 0$ . If  $\rho$  is the clockwise rotation of  $\pi/2$  around the origin  $(0, 0)$ , then the set  $H = \{\text{id}_{S^1}, \rho, \rho^2, \rho^3\}$  is a permutant for  $G$ .

**Example 2.14.** Let us consider the set  $X$  of the vertices of a cube in  $\mathbb{R}^3$ , and assume that  $\Phi$  is the set of all functions from  $X$  to  $[0, 1]$ . Let  $G$  be the group of the orientation-preserving isometries of  $\mathbb{R}^3$

that take  $X$  to  $X$ . Let  $\pi_1, \pi_2, \pi_3$  be the three planes that contain the center of mass of  $X$  and are parallel to a face of the cube. Let  $h_i: X \rightarrow X$  be the orthogonal symmetry with respect to  $\pi_i$ , for  $i \in \{1, 2, 3\}$ . We have that the set  $H = \{h_1, h_2, h_3\}$  is an orbit under the conjugation action of  $G$  on  $\text{Aut}_\Phi(X)$ , and therefore a permutant for  $G$ .

**Remark 2.15.** If the group  $G$  is Abelian, every finite subset of  $G$  is a permutant for  $G$ , since the conjugation action of  $G$  on  $\text{Aut}_\Phi(X)$  is just the identity.

**Remark 2.16.** In this section, the symbol  $\|\cdot\|_\infty$  has been used to denote the max-norm of functions. With a slight abuse of notation, in the rest of the paper such a symbol will be also used to denote the max-norm of points of  $\mathbb{R}^m$ , i.e.,  $\|(\alpha_1, \dots, \alpha_m)\|_\infty := \max_{1 \leq i \leq m} |\alpha_i|$ .

### 3. BUILDING GEOS FROM SYMMETRIC FUNCTIONS

**Definition 3.1.** Let  $C$  be a symmetric subset of  $\mathbb{R}^n$ , i.e., a subset  $C$  such that  $\pi(C) = C$  for every permutation  $\pi$  of the coordinates. A function  $f: C \rightarrow \mathbb{R}$  is said to be **symmetric on  $C$**  if its value is the same no matter the order of its arguments. That is,

$$f(a_1, \dots, a_n) = f(a_{\pi(1)}, \dots, a_{\pi(n)})$$

for every  $(a_1, \dots, a_n) \in C$  and every permutation  $\pi$  of the set  $\{1, \dots, n\}$ .

**Proposition 3.2.** Let  $f$  be a continuous real-valued symmetric function defined on a compact symmetric subset  $K$  of  $\mathbb{R}^n$ . Then  $\bar{f}$  is the restriction of a continuous real-valued symmetric function  $\hat{f}$  defined on  $\mathbb{R}^n$ .

*Proof:* The Tietze Extension Theorem (Dugundji, 1966) implies that  $f$  can be extended to a continuous function  $\hat{f}: \mathbb{R}^n \rightarrow \mathbb{R}$ . If  $S_n$  is the symmetric group over the set  $\{1, \dots, n\}$ , we can easily check that the function  $\bar{f}(a_1, \dots, a_n) := \frac{1}{n!} \sum_{\pi \in S_n} \hat{f}(a_{\pi(1)}, \dots, a_{\pi(n)})$  has the wanted property.

Proposition 3.2 guarantees that the concept of continuous real-valued symmetric function defined on a compact symmetric subset  $K$  of  $\mathbb{R}^n$  coincides with the concept of restriction to  $K$  of a continuous real-valued symmetric function defined on  $\mathbb{R}^n$ .

Let  $\mathcal{S}: \mathbb{R}^n \rightarrow \mathbb{R}$  be a symmetric function. If  $H = \{h_i\}_{i=1}^n$  is a non-empty permutant for  $G \subseteq \text{Aut}_\Phi(X)$ , then we can define an operator  $\mathcal{S}_H: \Phi \rightarrow \mathbb{R}_b^X$  by setting, for any  $\varphi \in \Phi$ ,

$$\mathcal{S}_H(\varphi) := \mathcal{S}(\varphi \circ h_1, \dots, \varphi \circ h_n),$$

where  $\mathcal{S}(\varphi \circ h_1, \dots, \varphi \circ h_n)(x) := \mathcal{S}((\varphi \circ h_1)(x), \dots, (\varphi \circ h_n)(x))$  for every  $x \in X$ .

**Proposition 3.3.** If  $\mathcal{S}: \mathbb{R}^n \rightarrow \mathbb{R}$  is a symmetric function and  $G \subseteq \text{Aut}_\Phi(X)$ , then  $\mathcal{S}_H$  is a GEO from  $\Phi$  to  $\mathbb{R}_b^X$  with respect to the identity homomorphism  $\text{id}_G: G \rightarrow G$ .

*Proof:* For every  $g \in G$ , it holds that

$$\begin{aligned} \mathcal{S}_H(\varphi \circ g) &= \mathcal{S}(\varphi \circ g \circ h_1, \dots, \varphi \circ g \circ h_n) \\ &= \mathcal{S}((\varphi \circ h_{\pi_g(1)}) \circ g, \dots, (\varphi \circ h_{\pi_g(n)}) \circ g) \\ &= \mathcal{S}((\varphi \circ h_1) \circ g, \dots, (\varphi \circ h_n) \circ g) \\ &= \mathcal{S}(\varphi \circ h_1, \dots, \varphi \circ h_n) \circ g \\ &= \mathcal{S}_H(\varphi) \circ g, \end{aligned}$$

where  $\pi_g$  is the permutation such that  $g \circ h_i \circ g^{-1} = h_{\pi_g(i)}$ , that is  $g \circ h_i = h_{\pi_g(i)} \circ g$ . Therefore,  $\mathcal{S}_H(\varphi \circ g) = \mathcal{S}_H(\varphi) \circ g$ , and hence  $\mathcal{S}_H$  is a GEO.

**Corollary 3.4.** If  $\mathcal{S}: \mathbb{R}^n \rightarrow \mathbb{R}$  is a symmetric function and its restriction to  $\text{Im}(\Phi)^n$  is non-expansive, then  $\mathcal{S}_H$  is a GENEIO from  $\Phi$  to  $\mathbb{R}_b^X$  with respect to  $\text{id}_G$ .

*Proof:* It is sufficient to prove the non-expansivity of  $\mathcal{S}_H$  with respect to the max-norms on  $\Phi$  and  $\mathbb{R}_b^X$ , since the group equivariance is already granted by Proposition 3.3. If  $\varphi, \psi \in \Phi$ , then for every  $x \in X$

$$\begin{aligned} &|(\mathcal{S}_H(\varphi))(x) - (\mathcal{S}_H(\psi))(x)| \\ &= |\mathcal{S}(\varphi \circ h_1, \dots, \varphi \circ h_n)(x) - \mathcal{S}(\psi \circ h_1, \dots, \psi \circ h_n)(x)| \\ &\leq \|((\varphi \circ h_1)(x) - (\psi \circ h_1)(x), \dots, (\varphi \circ h_n)(x) - (\psi \circ h_n)(x))\|_\infty \\ &= \max_{1 \leq i \leq n} |(\varphi \circ h_i)(x) - (\psi \circ h_i)(x)| \\ &\leq \max_{1 \leq i \leq n} \|\varphi \circ h_i - \psi \circ h_i\|_\infty \\ &= \|\varphi - \psi\|_\infty. \end{aligned}$$

In conclusion,  $\|\mathcal{S}_H(\varphi) - \mathcal{S}_H(\psi)\|_\infty \leq \|\varphi - \psi\|_\infty$  and  $\mathcal{S}_H$  is a GENEIO.

So far we have shown how to construct GEOs associated with a symmetric function. These operators are actually GENEIOs if the function they are associated with is non-expansive. In the next sections, we will show how to adapt this concept to build GENEIOs even in the presence of symmetric functions that are not non-expansive.

We stress that our approach requires no integration over the (possibly infinite and large) group  $G$ , but just the availability of a permutant and the computation of a symmetric function. This approach generalizes the method introduced in Camporesi et al. (2018), concerning the symmetric function  $\mathcal{S}(a_1, \dots, a_n) = \frac{1}{n} \sum_{i=1}^n a_i$ .

### 4. APPROXIMATING SYMMETRIC FUNCTIONS WITH SYMMETRIC POLYNOMIALS

Let us now explore the concept of approximation of symmetric functions by symmetric polynomials. For more details, we refer the reader to Davidson and Donsig (2009), Blum-Smith and Coskey (2017). In the sequel, we will denote the symmetric

group over the set  $\{1, \dots, n\}$  as  $S_n$ . Let  $K$  be a compact metric space, and  $\mathcal{C}(K)$  be the vector space of continuous real-valued functions on  $K$ . With a slight abuse of notation, in the following we will confuse each polynomial with the function it represents, restricted to the domain we are considering. Furthermore, if  $I$  is a finite subset of  $\mathbb{N}^n$ , we will say that a polynomial  $\sum_{(k_1, \dots, k_n) \in I} c_{k_1, \dots, k_n} y_1^{k_1} \cdot \dots \cdot y_n^{k_n}$  is **symmetric** if  $\pi(I) = I$ , and  $c_{k_1, \dots, k_n} = c_{\pi(k_1), \dots, \pi(k_n)}$  for every multi-index  $(k_1, \dots, k_n) \in I$  and every permutation  $\pi \in S_n$ .

**Definition 4.1.** Davidson and Donsig (2009) A subset  $A$  of  $\mathcal{C}(K)$  is an **algebra** if it is a vector subspace of  $\mathcal{C}(K)$  that is closed under multiplication (i.e., if  $f, g \in A$  then  $f \cdot g \in A$ ). A set  $S$  of functions on  $K$  **separates points** if for each pair of points  $s, t \in K$  there is a function  $f \in S$  such that  $f(s) \neq f(t)$ . A set  $S$  of functions on  $K$  **vanishes** at  $s \in K$  if  $f(s) = 0$  for all  $f \in S$ .

**Theorem 4.2 (Stone - Weierstrass Theorem).** Davidson and Donsig (2009) An algebra  $A$  of continuous real-valued functions on a compact metric space  $K$  that separates points and does not vanish at any point is dense in  $\mathcal{C}(K)$  with respect to the max-norm referred to the domain  $K$ .

**Corollary 4.3.** Davidson and Donsig (2009) Let  $K$  be a compact subset of  $\mathbb{R}^n$ . The algebra of all polynomials  $p(y_1, \dots, y_n)$  in  $n$  variables is dense in  $\mathcal{C}(K)$  with respect to the max-norm referred to the domain  $K$ .

This theorem allows us to approximate a continuous symmetric function  $\mathcal{S}: K \rightarrow \mathbb{R}$  by a polynomial with arbitrary accuracy, provided that  $K$  is a compact subset of  $\mathbb{R}^n$ . However, this is not exactly what we need, as we would like such a polynomial to be symmetric. This can be obtained by a symmetrization of the previously found polynomial, as shown by the next proposition, which proves that any symmetric continuous function on a compact and symmetric domain can be approximated with arbitrary precision by a symmetric polynomial.

**Proposition 4.4.** Let  $K$  be a compact subset of  $\mathbb{R}^n$ , verifying the property  $\pi(K) = K$  for every  $\pi \in S_n$ . If  $\mathcal{S}|_K: K \rightarrow \mathbb{R}$  is the restriction to  $K$  of a continuous symmetric function  $\mathcal{S}: \mathbb{R}^n \rightarrow \mathbb{R}$  and  $\|\cdot\|_\infty$  is the max-norm referred to the domain  $K$ , then for every  $\varepsilon > 0$  there exists a symmetric polynomial  $q$  in  $n$  variables such that  $\|\mathcal{S}|_K - q|_K\|_\infty \leq \varepsilon$ .

*Proof:* From Corollary 4.3 it follows that there exists a polynomial  $p: \mathbb{R}^n \rightarrow \mathbb{R}$  such that  $\|\mathcal{S}|_K - p|_K\|_\infty \leq \varepsilon$ . Let us now define the symmetric polynomial  $q(a_1, \dots, a_n): = \frac{1}{n!} \sum_{\pi \in S_n} p(a_{\pi(1)}, \dots, a_{\pi(n)})$ . If  $a = (a_1, \dots, a_n) \in K$ , we define  $a_\pi = (a_{\pi(1)}, \dots, a_{\pi(n)})$  for every permutation  $\pi \in S_n$ . Then

$$\begin{aligned} \|\mathcal{S}|_K - q|_K\|_\infty &= \max_{a \in K} |\mathcal{S}(a) - q(a)| \\ &= \max_{a \in K} \left| \mathcal{S}(a) - \frac{1}{n!} \sum_{\pi \in S_n} p(a_\pi) \right| \\ &= \max_{a \in K} \left| \frac{1}{n!} \sum_{\pi \in S_n} \mathcal{S}(a) - \frac{1}{n!} \sum_{\pi \in S_n} p(a_\pi) \right| \end{aligned}$$

$$\begin{aligned} &\leq \frac{1}{n!} \max_{a \in K} \sum_{\pi \in S_n} |\mathcal{S}(a) - p(a_\pi)| \\ &= \frac{1}{n!} \max_{a \in K} \sum_{\pi \in S_n} |\mathcal{S}(a_\pi) - p(a_\pi)| \\ &= \frac{1}{n!} \sum_{\pi \in S_n} \max_{a \in K} |\mathcal{S}(a_\pi) - p(a_\pi)| \\ &= \frac{1}{n!} \sum_{\pi \in S_n} \max_{a \in K} |\mathcal{S}(a) - p(a)| \\ &= \frac{1}{n!} \sum_{\pi \in S_n} \|\mathcal{S}|_K - p|_K\|_\infty \\ &\leq \frac{1}{n!} \sum_{\pi \in S_n} \varepsilon = \varepsilon. \end{aligned}$$

**Definition 4.5.** The **elementary symmetric polynomials** in the  $n$  variables  $a_1, \dots, a_n$ , also called **elementary symmetric functions**, are defined as:

$$\begin{aligned} \sigma_1 &:= a_1 + \dots + a_n \\ \sigma_2 &:= a_1 \cdot a_2 + a_1 \cdot a_3 + \dots + a_{n-1} \cdot a_n = \sum_{1 \leq i < j \leq n} a_i \cdot a_j \\ &\vdots \\ \sigma_r &:= \sum_{1 \leq i_1 < i_2 < \dots < i_r \leq n} a_{i_1} \cdot a_{i_2} \cdot \dots \cdot a_{i_r} = \sum_{1 \leq i_1 < i_2 < \dots < i_r \leq n} \prod_{j=1}^r a_{i_j} \\ &\vdots \\ \sigma_n &:= a_1 \cdot a_2 \cdot \dots \cdot a_n. \end{aligned}$$

We now recall an important result in the theory of symmetric polynomials:

**Theorem 4.6. (Fundamental Theorem on Symmetric Polynomials).** Blum-Smith and Coskey (2017) Any symmetric polynomial in  $n$  variables  $a_1, \dots, a_n$  is representable in a unique way as a polynomial in the elementary symmetric polynomials  $\sigma_1, \dots, \sigma_n$ .

**Remark 4.7.** It is a well know fact (see Rao, 2005; Davidson and Donsig, 2009; Blum-Smith and Coskey, 2017) that the proofs of Theorems 4.2 and 4.6 are constructive. This means that, if  $K$  is a compact symmetric subset of  $\mathbb{R}^n$ , and the restriction  $\mathcal{S}|_K$  of a continuous symmetric function  $\mathcal{S}: \mathbb{R}^n \rightarrow \mathbb{R}$  is given, we are able to effectively approximate  $\mathcal{S}|_K$  with an error less than  $\varepsilon$  by the restriction to  $K$  of an explicitly defined polynomial in the elementary symmetric functions.

In conclusion, if an equivariance group  $G$  is chosen and a GEO  $F$  is built by applying Proposition 3.3 to the continuous symmetric function  $\mathcal{S}: \mathbb{R}^n \rightarrow \mathbb{R}$ , we can approximate  $F$  in the following way, provided that  $X$  and  $\Phi$  are compact. First of all, we can approximate the continuous function  $\mathcal{S}$  by a polynomial  $p: \mathbb{R}^n \rightarrow \mathbb{R}$ , with an arbitrarily small error  $\varepsilon$  on the symmetric set  $\text{Im}(\Phi)^n$ , which is guaranteed to be compact

by Proposition 2.2. Then, we can consider the symmetric polynomial  $q(a_1, \dots, a_n) := \frac{1}{n!} \sum_{\pi \in S_n} p(a_{\pi(1)}, \dots, a_{\pi(n)})$ . Finally, we can consider the GEO  $F'$  defined by setting  $F'(\varphi) := q(\varphi \circ h_1, \dots, \varphi \circ h_n)$  for every  $\varphi \in \Phi$ . Since  $H \subseteq \text{Aut}_\Phi(X)$ ,  $\|F(\varphi) - F'(\varphi)\|_\infty = \max_{x \in X} |\mathcal{S}(\varphi(h_1(x)), \dots, \varphi(h_n(x))) - q(\varphi(h_1(x)), \dots, \varphi(h_n(x)))| \leq \|S_{\text{Im}(\Phi)^n} - q_{\text{Im}(\Phi)^n}\|_\infty \leq \varepsilon$  for any  $\varphi \in \Phi$ , and hence the operator  $F'$  can be chosen arbitrarily close to  $F$ .

### 5. BUILDING GENEOS FROM POLYNOMIALS IN THE ELEMENTARY SYMMETRIC FUNCTIONS

Proposition 2.2 shows that  $\text{Im}(\Phi)$  is compact. Moreover, the equality  $\pi(\text{Im}(\Phi)^n) = \text{Im}(\Phi)^n$  trivially holds for every  $\pi \in S_n$ . Therefore, Proposition 4.4 and the Fundamental Theorem on Symmetric Polynomials guarantee that the restriction to  $\text{Im}(\Phi)^n$  of any continuous symmetric functions can be approximated arbitrarily well by the restriction to  $\text{Im}(\Phi)^n$  of a polynomial in the elementary symmetric functions, defined as

$$\tilde{S}(a_1, \dots, a_n) = \sum_{k_1=0}^{m_1} \cdots \sum_{k_n=0}^{m_n} c_{k_1, \dots, k_n} \prod_{i=1}^n \sigma_i^{k_i}(a_1, \dots, a_n),$$

where  $m_i \in \mathbb{N}$  for every  $i \in \{1, \dots, n\}$ ,  $c_{k_1, \dots, k_n} \in \mathbb{R}$  for every  $k_1 \in \{0, \dots, m_1\}, \dots, k_n \in \{0, \dots, m_n\}$  and  $\sigma_i$  is the  $i$ -th elementary symmetric polynomial for every  $i \in \{1, \dots, n\}$ . From Proposition 3.3, we already know that the associated operator is a GEO. We can indeed obtain a GENEIO by applying Corollary 3.4 to a suitable multiple of  $\tilde{S}$ . In the sequel, we will need the following constants:

$$M_{\text{Im}(\Phi)^n} := \max_{\alpha \in \text{Im}(\Phi)^n} \|\alpha\|_\infty = \max_{\varphi \in \Phi} \|\varphi\|_\infty \tag{5.1}$$

$$M_1 := \max_{1 \leq i \leq n} \left\{ k_i \binom{n}{i}^{k_i} i M_{\text{Im}(\Phi)^n}^{i k_i - 1} \right\} \tag{5.2}$$

$$M_2 := \max_{1 \leq i \leq n} \left\{ \binom{n}{i}^{k_i} M_{\text{Im}(\Phi)^n}^{i k_i} \right\}^{n-1} \tag{5.3}$$

$$C = n \sum_{k_1=0}^{m_1} \cdots \sum_{k_n=0}^{m_n} |c_{k_1, \dots, k_n}| M_1 M_2, \tag{5.4}$$

Let us consider a non-empty permutant  $H = \{h_i\}_{i=1}^n$  for  $G \subseteq \text{Aut}_\Phi(X)$ . We can define an operator  $\hat{S}_H: \Phi \rightarrow \mathbb{R}_b^X$  by setting

$$\hat{S}_H(\varphi) := \frac{1}{C} \tilde{S}(\varphi \circ h_1, \dots, \varphi \circ h_n)$$

for any  $\varphi \in \Phi$ , where  $\tilde{S}(\varphi \circ h_1, \dots, \varphi \circ h_n)(x) := \tilde{S}((\varphi \circ h_1)(x), \dots, (\varphi \circ h_n)(x))$  for every  $x \in X$  and  $C$  is the constant defined in (5.4).

**Theorem 5.1.** *If  $\tilde{S}$  is a polynomial in the  $n$  elementary symmetric functions, then  $\hat{S}_H$  is a GENEIO from  $\Phi$  to  $\mathbb{R}_b^X$  with respect to  $\text{id}_G$ .*

*Proof:* The thesis immediately follows from Corollary 3.4, once it is proved that the restriction of  $\frac{1}{C} \tilde{S}$  to  $\text{Im}(\Phi)^n$  is non-expansive. For every  $\alpha = (\alpha_1, \dots, \alpha_n), \beta = (\beta_1, \dots, \beta_n) \in \text{Im}(\Phi)^n$ , by applying Lemma 1.5 in **Appendix 1**, we have that

$$\begin{aligned} & \left| \frac{1}{C} \tilde{S}(\alpha) - \frac{1}{C} \tilde{S}(\beta) \right| \\ &= \left| \frac{1}{C} \sum_{k_1=0}^{m_1} \cdots \sum_{k_n=0}^{m_n} c_{k_1, \dots, k_n} \prod_{i=1}^n \sigma_i^{k_i}(\alpha) \right. \\ & \quad \left. - \frac{1}{C} \sum_{k_1=0}^{m_1} \cdots \sum_{k_n=0}^{m_n} c_{k_1, \dots, k_n} \prod_{i=1}^n \sigma_i^{k_i}(\beta) \right| \\ & \leq \frac{1}{C} \sum_{k_1=0}^{m_1} \cdots \sum_{k_n=0}^{m_n} |c_{k_1, \dots, k_n}| \left| \prod_{i=1}^n \sigma_i^{k_i}(\alpha) - \prod_{i=1}^n \sigma_i^{k_i}(\beta) \right| \\ & \leq \frac{1}{C} \sum_{k_1=0}^{m_1} \cdots \sum_{k_n=0}^{m_n} |c_{k_1, \dots, k_n}| n \|\alpha - \beta\|_\infty M_1 M_2 \\ & = \left( \frac{1}{C} n \sum_{k_1=0}^{m_1} \cdots \sum_{k_n=0}^{m_n} |c_{k_1, \dots, k_n}| M_1 M_2 \right) \|\alpha - \beta\|_\infty \\ & = \|\alpha - \beta\|_\infty. \end{aligned}$$

We have shown that  $\frac{1}{C} \tilde{S}$  is non-expansive and therefore the associated operator  $\hat{S}_H$  is a GENEIO.  $\square$

**Example 5.2.** *Let us consider the setting of Example 2.12 and the polynomial  $\tilde{S}(a_1, a_2) = -\sigma_1(a_1, a_2) + \sigma_2^2(a_1, a_2) - 3\sigma_1(a_1, a_2)\sigma_2(a_1, a_2) = -a_1 - a_2 + a_1^2 a_2^2 - 3a_1^2 a_2 - 3a_1 a_2^2$ . Then the operator*

$$\begin{aligned} \hat{S}_H(\varphi) &= \frac{1}{40} (-(\varphi \circ h) - (\varphi \circ h^{-1}) + (\varphi^2 \circ h) \cdot (\varphi^2 \circ h^{-1}) \\ & \quad - 3(\varphi^2 \circ h) \cdot (\varphi \circ h^{-1}) - 3(\varphi \circ h) \cdot (\varphi^2 \circ h^{-1})) \end{aligned}$$

is a GENEIO.

**Example 5.3.** *Let us consider the setting of Example 2.13 and the polynomial  $\tilde{S}(a_1, a_2, a_3, a_4) = \sigma_1(a_1, a_2, a_3, a_4) + \sigma_4(a_1, a_2, a_3, a_4) = a_1 + a_2 + a_3 + a_4 + a_1 a_2 a_3 a_4$ . In this case, the operator*

$$\begin{aligned} \hat{S}_H(\varphi) &= \frac{1}{1040} (\varphi + (\varphi \circ \rho) + (\varphi \circ \rho^2) + (\varphi \circ \rho^3) \\ & \quad + \varphi \cdot (\varphi \circ \rho) \cdot (\varphi \circ \rho^2) \cdot (\varphi \circ \rho^3)) \end{aligned}$$

is a GENEIO.

**Example 5.4.** *Let us consider the setting of Example 2.14 and the polynomial  $\tilde{S}(a_1, a_2, a_3) = \sigma_2(a_1, a_2, a_3) \cdot \sigma_3(a_1, a_2, a_3) = a_1^2 a_2^2 a_3 + a_1^2 a_2 a_3^2 + a_1 a_2^2 a_3^2$ . Then the operator*

$$\hat{S}_H(\varphi) = \frac{1}{162} ((\varphi^2 \circ h_1) \cdot (\varphi^2 \circ h_2) \cdot (\varphi \circ h_3))$$

$$\begin{aligned}
 &+ (\varphi^2 \circ h_1) \cdot (\varphi \circ h_2) \cdot (\varphi^2 \circ h_3) \\
 &+ (\varphi \circ h_1) \cdot (\varphi^2 \circ h_2) \cdot (\varphi^2 \circ h_3)
 \end{aligned}$$

is a GENEIO.

**Remark 5.5.** It is worth noticing that if we replace the constant  $C$ , defined as in (5.4), with any constant  $\bar{C} \geq C$ , the operator  $\hat{S}_H$  defined as in Theorem 5.1 is still a GENEIO. However, it should be specified that the larger the constant  $\bar{C}$  is, the more difficult it becomes to distinguish different signals, since their distance is smaller. For this reason, a larger  $\bar{C}$  means that more information is lost when applying the GENEIO. Nonetheless, in some cases rough constants which are easier to compute may be preferred. For this reason, we present here a possible alternative to  $C$ . Let  $c = \max |c_{k_1, \dots, k_n}|$ ,  $m = \max_j m_j$ ,  $\mu = \prod_{j=1}^n (m_j + 1)$  and  $\bar{M}_{\text{Im}(\Phi)^n} = \max \{M_{\text{Im}(\Phi)^n}, 1\}$ , where  $M_{\text{Im}(\Phi)^n}$  is defined as in (5.1). We define the constant

$$\bar{C} = c\mu n^2 m \binom{n}{\lfloor \frac{n}{2} \rfloor}^{mn} \bar{M}_{\text{Im}(\Phi)^n}^{mn^2-1}. \tag{5.5}$$

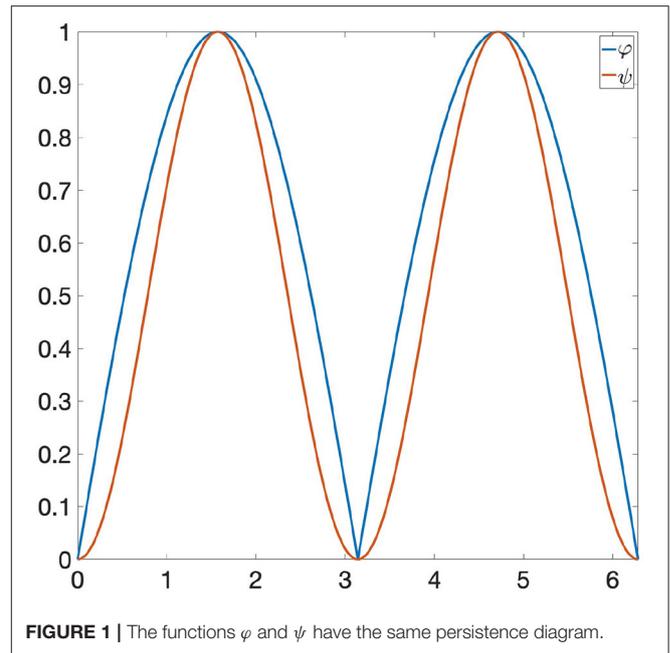
We can easily show that  $\bar{C} \geq C$ . Therefore, the operator  $\tilde{S}_H$ , obtained by replacing  $C$  with  $\bar{C}$  in the definition of  $\hat{S}_H$  and applying Theorem 5.1, is still a GENEIO. In Example 5.2,  $\bar{C} = 2,304$ , which is far larger than  $C = 40$ . In Example 5.3,  $\bar{C} = 82,944$ , while  $C = 1,040$ . Finally, in Example 5.4  $\bar{C} = 972$  and  $C = 162$ . We stress that the constant  $C$  defined in (5.4) is optimal if we do not add any further assumption. We can realize this by applying Theorem 5.1 to the symmetric function  $\tilde{S}(a_1) = \sigma_1(a_1) = a_1$  in just one variable, provided that  $\Phi$  is the collection of all non-expansive functions from  $X := [0, 1]$  to itself, and we set both  $G$  and the permutant  $H$  equal to the trivial group containing only the identity of  $X$ . On the one hand, it can be easily checked that in this case the GEO  $\tilde{S}_H$  defined by applying Proposition 3.3 is the identity map, and hence a GENEIO. Therefore, any constant  $C'$  that could replace  $C$  in the definition of  $\hat{S}_H$  must be not smaller than 1 in order to preserve the non-expansivity of  $\hat{S}_H$ . On the other hand, we can immediately see that  $C = M_1 = M_2 = M_{\text{Im}(\Phi)^n} = 1$ . It follows that  $C' \geq C$ .

## 6. GENEOS INCREASE OUR ABILITY TO DISTINGUISH DATA

In this section, we illustrate a few examples showing that our approach can produce useful non-linear GENEOS and increase our ability to distinguish data. As already discussed in the Introduction, it is of fundamental importance to make a large number of GENEOS available in machine learning, as each of them models a data-observer pair. The results presented in this paper could be of great help in the task of extending the set of available GENEOS and the consequent possibilities of using them in applications.

The next example shows that our new method indeed extends the approach introduced in Botteghi et al. (2020).

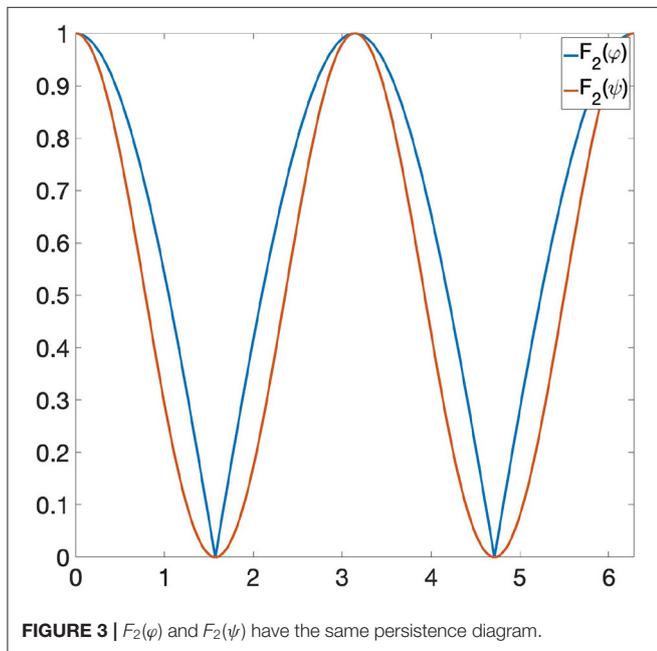
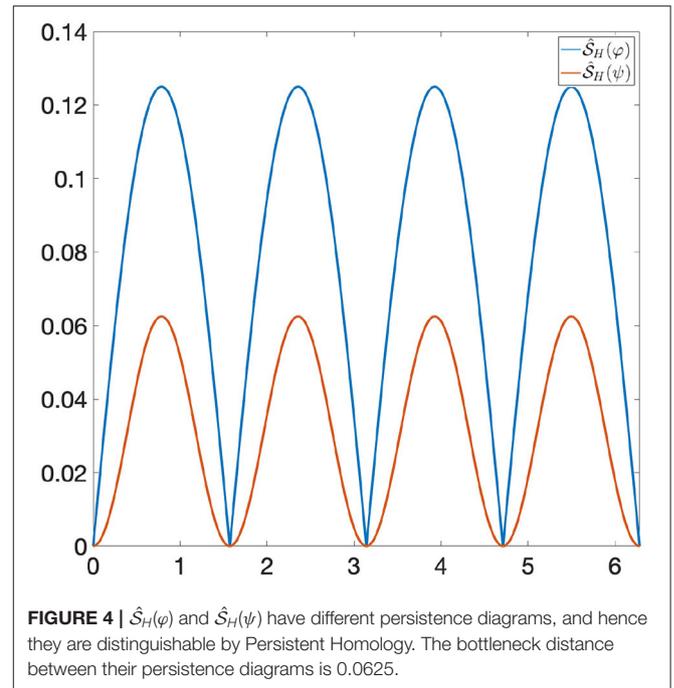
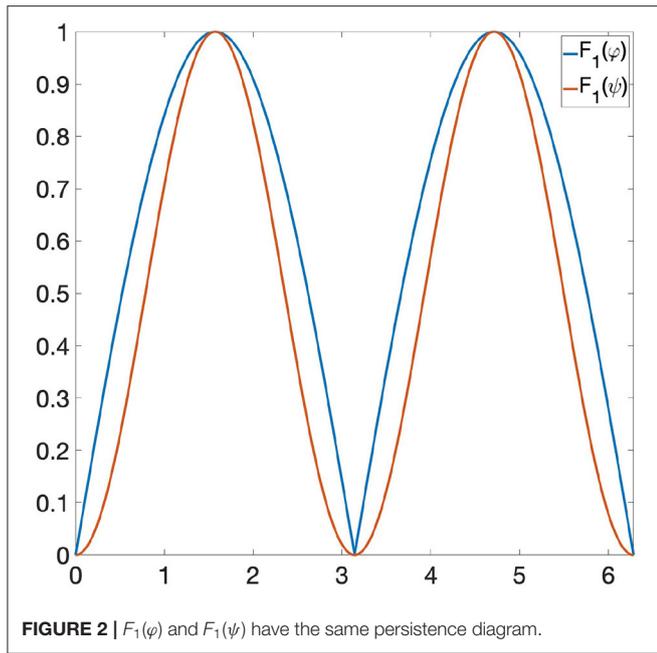
**Example 6.1.** Let us set  $X = \{1, 2, 3\}$  and  $\Phi$  equal to the set of all functions from  $X$  to  $[0, 1]$ . In this case,  $\text{Aut}_\Phi(X) = S_3$ ,



and we define  $G = S_3$ . We now consider the symmetric function  $\tilde{S}(a_1, a_2) = a_1 \cdot a_2 = \sigma_2(a_1, a_2)$  and the permutant  $H = \{(1, 2, 3), (1, 3, 2)\} =: \{h_1, h_2\}$ . From Theorem 5.1 we get that  $\hat{S}_H(\varphi) = \frac{(\varphi \circ h_1) \cdot (\varphi \circ h_2)}{4}$  is a GENEIO. Such an operator is not linear, and hence it cannot be obtained by the method described in Botteghi et al. (2020).

In next Example 6.2 we illustrate the synergy between GENEOS and TDA. TDA is mainly grounded on Persistent Homology (PH), which is an algebraic topological theory devised to describe “holes” in geometrical data, focusing on their persistence under the action of noise. In particular, TDA takes benefit from the topological comparison of data by means of persistence diagrams, which are the main tools in PH. For more details about TDA and PH, we refer the interested reader to Edelsbrunner and Harer (2008), Edelsbrunner and Morozov (2013). Example 6.2 shows that the use of new GENEOS increases our ability of distinguishing data by persistence diagrams.

**Example 6.2.** Let us consider the following functions:  $\varphi(x) = |\sin(x)|$  and  $\psi(x) = \sin(x)^2 \in \Phi$ , where  $\Phi$  is the space of all 1-Lipschitz functions from the unit circle  $S^1$  to  $[0, 1]$  and the invariance group  $G$  is composed of all rotations of  $S^1$ . If we are looking at  $\Phi$  only through Persistent Homology, then  $\varphi$  and  $\psi$  are indistinguishable, since they induce the same persistence diagram (see Figure 1). Let us now consider the GENEOS  $F_1 = \text{id} : \Phi \rightarrow \Phi$  and  $F_2(\varphi) = \varphi \circ \rho_{\frac{\pi}{2}}$ , where  $\rho_{\frac{\pi}{2}}$  is the clockwise rotation through a  $\frac{\pi}{2}$  angle. We highlight the fact that the joint use of TDA,  $F_1$  and  $F_2$  does not allow us to distinguish  $\varphi$  from  $\psi$ , since the functions  $F_1(\varphi)$  and  $F_1(\psi)$  have the same persistence diagram, and the same happens for  $F_2(\varphi)$  and  $F_2(\psi)$  (see Figures 2, 3). However, if we consider the symmetric function  $\sigma_2$  and the permutant  $H = \{\text{id}_{S^1}, \rho_{\frac{\pi}{2}}\}$ , from Theorem 5.1 we get that  $\hat{S}_H(\varphi) =$



In Example 6.3 we show how the pseudo-metrics associated with non-linear GENEOS are much more flexible than those generated by linear operators, thus guaranteeing a wider range of applications.

**Example 6.3.** Let us consider  $\Phi$  as the set of all functions from the set with just two elements  $X = \{A, B\}$  to  $[0, 1]$ . The functions in  $\Phi$  can be described by ordered pairs  $(\varphi(A), \varphi(B))$ . In this setting, the group  $G$  is composed of the permutations of two elements. As usual, in  $\Phi$  we have the metric  $D_\Phi(\varphi_1, \varphi_2) = \|\varphi_1 - \varphi_2\|_\infty$ , therefore the distance between the functions  $(0, 0), (0, 1), (1, 0), (1, 1)$  is always 1. Suppose that we need a GENEIO  $F$  such that the pseudo-metric  $\|F(\varphi_1) - F(\varphi_2)\|_\infty$  vanishes between functions with a null component, while maintaining positive the distance between  $(1, 1)$  and the other three functions belonging to  $\Phi$ . No linear GENEIO can induce such a pseudo-metric, since if a linear transformation maps  $(1, 0)$  and  $(0, 1)$  to  $(0, 0)$ , it must also map  $(1, 1)$  to  $(0, 0)$ . It is worth noticing that through the GENEIO associated with the elementary symmetric function  $\sigma_2(a_1, a_2) = a_1 \cdot a_2$  and with the permutant  $H = G$ , we can obtain a pseudo-metric with the desired property.

$\frac{1}{4}(\varphi \cdot (\varphi \circ \rho_{\frac{\pi}{2}}))$  is a non-linear GENEIO. We observe that the joint use of TDA and  $\hat{S}_H$  allows us to distinguish  $\varphi$  and  $\psi$ , since the persistence diagrams of the functions  $\hat{S}_H(\varphi)$  and  $\hat{S}_H(\psi)$  are different from each other (see **Figure 4**).

## 7. CONCLUSIONS

Finally, we stress another important aspect of the GENEOS. The use of GENEOS can be seen as a methodology for changing the max-norm metric that we use to compare functions in  $\Phi$ . We can indeed change the distance  $\|\varphi_1 - \varphi_2\|_\infty$  into the pseudo-metric  $\|F(\varphi_1) - F(\varphi_2)\|_\infty$ , where  $F$  is a GENEIO.

In our paper, we have introduced a new method to build GENEOS, grounded on the concepts of symmetric function and permutant. Our main goal is the one of building a good theory of GENEOS, making available methods to define and use these operators in machine learning. The main advantage of our approach is the fact that it requires no integration over the (possibly infinite and large) group  $G$ , but just the availability of a permutant and the computation of a symmetric function. Many

lines of research still remain to be explored in this field. For example, the reader can observe that in section 3 the requirement that the function  $\mathcal{S}$  is symmetric could be weakened without losing the property that  $\mathcal{S}_H$  is a GEO. It would be indeed sufficient to assume that  $\mathcal{S}$  is invariant when we apply to its argument any permutation corresponding to the permutation of  $H$  associated with the conjugation action  $h \mapsto g \circ h \circ g^{-1}$  that is defined by any  $g \in G$ . We have decided to postpone the research concerning this extension of the theory, since it would have added a dependence on  $H$  of the choice of  $\mathcal{S}$ , thereby introducing some technicalities. Another line of research that we would like to explore in the future is the possibility of adapting our approach to permutant measures, which are a sort of extension of the concept of permutant to the case that  $H$  has infinite cardinality. However, the most challenging problem we will have to face is likely to be the proof or disproof of the natural conjecture that each non-linear  $G$ -equivariant non-expansive operator can be produced (or at least well approximated) by applying our new technique to suitable symmetric functions and permutants, provided that the group  $G$  transitively acts on a finite signal domain. We plan to devote some subsequent papers to these topics.

## REFERENCES

- Anselmi, F., Evangelopoulos, G., Rosasco, L., and Poggio, T. (2019). Symmetry-adapted representation learning. *Pattern Recogn.* 86, 201–208. doi: 10.1016/j.patcog.2018.07.025
- Anselmi, F., Rosasco, L., and Poggio, T. (2016). On invariance and selectivity in representation learning. *Inform. Inference J. IMA* 5, 134–158. doi: 10.1093/imaiai/iaw009
- Bengio, Y., Courville, A., and Vincent, P. (2013). Representation learning: a review and new perspectives. *IEEE Trans. Pattern Anal. Mach. Intell.* 35, 1798–1828. doi: 10.1109/TPAMI.2013.50
- Bergomi, M. G., Frosini, P., Giorgi, D., and Quercioli, N. (2019). Towards a topological-geometrical theory of group equivariant non-expansive operators for data analysis and machine learning. *Nat. Mach. Intell.* 1, 423–433. doi: 10.1038/s42256-019-0087-3
- Blum-Smith, B., and Coskey, S. (2017). The fundamental theorem on symmetric polynomials: history's first whiff of Galois theory. *College Math. J.* 48, 18–29. doi: 10.4169/college.math.j.48.1.18
- Botteghi, S., Brasini, M., Frosini, P., and Quercioli, N. (2020). On the finite representation of group equivariant operators via permutant measures. *arXiv preprint arXiv:2008.06340*.
- Camporesi, F., Frosini, P., and Quercioli, N. (2018). "On a new method to build group equivariant operators by means of permutants," in *2nd International Cross-Domain Conference for Machine Learning and Knowledge Extraction (CD-MAKE)*, eds A. Holzinger, P. Kieseberg, A. M. Tjoa, and E. Weippl (Hamburg: Springer International Publishing), 265–272. doi: 10.1007/978-3-319-99740-7\_18
- Carrieri, A. P., Haiminen, N., Maudsley-Barton, S., Gardiner, L.-J., Murphy, B., Mayes, A. E., et al. (2021). Explainable AI reveals changes in skin microbiome composition linked to phenotypic differences. *Sci. Rep.* 11:4565. doi: 10.1038/s41598-021-83922-6
- Chachólski, W., Gregorio, A. D., Quercioli, N., and Tombari, F. (2020). Landscapes of data sets and functoriality of persistent homology. *arXiv preprint arXiv:2002.05972*.
- Cohen, T., and Welling, M. (2016). "Group equivariant convolutional networks," in *International Conference on Machine Learning* (New York, NY: ICML), 2990–2999.
- Davidson, K., and Donsig, A. (2009). *Real Analysis and Applications: Theory in Practice. Undergraduate Texts in Mathematics*. New York, NY: Springer. doi: 10.1007/978-0-387-98098-0
- Dugundji, J. (1966). *Topology*. Boston, MA: Allyn and Bacon.
- Edelsbrunner, H., and Harer, J. (2008). "Persistent homology—a survey," in *Surveys on Discrete and Computational Geometry*, eds J. E. Goodman, J. Pach, and R. Pollack (Providence, RI: American Mathematical Society), 257–282. doi: 10.1090/conm/453/08802
- Edelsbrunner, H., and Morozov, D. (2013). "Persistent homology: theory and practice," in *European Congress of Mathematics* (Zürich: European Mathematical Society), 31–50. doi: 10.4171/120-1/3
- Frosini, P. (2009). Does intelligence imply contradiction? *Cogn. Syst. Res.* 10, 297–315. doi: 10.1016/j.cogsys.2007.07.009
- Frosini, P. (2016). "Towards an observer-oriented theory of shape comparison: position paper," in *Proceedings of the Eurographics 2016 Workshop on 3D Object Retrieval, 3DOR '16* (Goslar: Eurographics Association), 5–8.
- Frosini, P., and Jabłoński, G. (2016). Combining persistent homology and invariance groups for shape comparison. *Discrete Comput. Geom.* 55, 373–409. doi: 10.1007/s00454-016-9761-y
- Frosini, P., and Quercioli, N. (2017). "Some remarks on the algebraic properties of group invariant operators in persistent homology," in *1st International Cross-Domain Conference for Machine Learning and Knowledge Extraction (CD-MAKE)*, eds A. Holzinger, P. Kieseberg, A. M. Tjoa, and E. Weippl (Reggio: Springer International Publishing), 14–24. doi: 10.1007/978-3-319-66808-6\_2
- Hicks, S. A., Isaksen, J. L., Thambawita, V., Ghouse, J., Ahlberg, G., Linneberg, A., et al. (2021). Explaining deep neural networks for knowledge discovery in electrocardiogram analysis. *Sci. Rep.* 11:10949. doi: 10.1038/s41598-021-90285-5
- Mallat, S. (2012). Group invariant scattering. *Commun. Pure Appl. Math.* 65, 1331–1398. doi: 10.1002/cpa.21413
- Mallat, S. (2016). Understanding deep convolutional networks. *Philos. Trans. R. Soc A Math. Phys. Eng. Sci.* 374:20150203. doi: 10.1098/rsta.2015.0203
- Rao, N. V. (2005). The Stone-Weierstrass theorem revisited. *Am. Math. Monthly* 112, 726–729. doi: 10.1080/00029890.2005.11920244
- Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat. Mach. Intell.* 1, 206–215. doi: 10.1038/s42256-019-0048-x

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

## AUTHOR CONTRIBUTIONS

PF devised the project. All authors contributed to the manuscript. All authors read and approved the final manuscript.

## ACKNOWLEDGMENTS

This research has been partially supported by INdAM-GNSAGA. FC thanks Davide Moroni and Maria Antonietta Pascali for their helpful advise and support.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/frai.2022.786091/full#supplementary-material>

- Worrall, D. E., Garbin, S. J., Turmukhambetov, D., and Brostow, G. J. (2017). "Harmonic networks: Deep translation and rotation equivariance," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)* (Honolulu, HI: IEEE), 7168–7177. doi: 10.1109/CVPR.2017.758
- Zhang, C., Voinea, S., Evangelopoulos, G., Rosasco, L., and Poggio, T. (2015). "Discriminative template learning in group-convolutional networks for invariant speech representations," in *Interspeech-2015* (Dresden: International Speech Communication Association), 3229–3233. doi: 10.21437/Interspeech.2015-650

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

*Copyright © 2022 Conti, Frosini and Quercioli. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*